catheterization in order to help characterize left ventricular function. Attempts to determine ventricular volume using one or two dimensional information from ultrasound has not demonstrated the accuracy of angiography. Therefore, three-dimensional information should provide a more accurate means of non-invasively assessing the state of the left ventricle.

C. Highlights of Research Progress

During the past year we have evaluated a large portion of the first phase of our implementation. As mentioned in previous reports the first phase involves the manual outlining of scans with a light pen and the interpolation of an ad hoc model to the data. The system as it stands now is entirely data directed and hence has no notion of what kind of organ it is looking at. However we expect that it will still give more accurate volumes than any now obtainable, and thus will provide the medical impetus for continued research into model directed approaches. The following is a list of accomplishments over the past year:

1. Elimination of most remaining bugs in the data acquisition and SUMEX modelling software.

2. Development of a user's guide, to the point where the system is now routinely used by a person other than the system developer.

3. An evaluation of the ability of the system to locate a point in three dimensions, the result of which shows that our accuracy is about 5-6mm. The sources of this uncertainty were evaluated separately and found to come about equally from the resolution of the ultrasound scanner (2mm), the resolution of the position locator (2mm) and the combination light pen and display monitor. An additional error source, namely the exact parameters describing the mounting of the 3D position locating device to the ultrasound scanner, was essentially eliminated by employing a relaxation procedure to optimize these parameters.

4. The system's ability to calculate balloon volumes was tested on 10 balloons ranging between 257 and 2092 cc. Each balloon was imaged 3 times. Most of the errors between measured and calculated volumes fell within 2%.

5. The system's ability to calculate kidney volumes was tested on two each of pork, sheep and human kidneys, again imaging each organ 3 times. Most of the errors fell within 5%.

6. A clinical study of the ability of the system to predict fetal weight has been underway for about 9 months. A clinical protocol, which directs how scans are to be acquired from the patients, has been developed and refined over that time. This protocol has evolved into a "bottom up" method of collecting on video tape and floppy disk essentially all the information that it is possible to collect from the fetus. Although all this information is not currently being used it will remain

available for use by later model directed versions of the system. In the present study we are comparing the ability of fetal head and trunk 3D reconstructions versus a series of 3D points and axes to predict fetal weight. Two patient populations are being tested: dead babies in a water tank (to test the accuracy in an "ideal" case), and live term babies in utero. In each case calculated volumes are compared against measured weights in a multiple regression using the program STATPACK, which is present on SUMEX. At present we have examined 26 dead babies and 31 live babies. The results so far indicate, somewhat surprisingly, that a combination of simple measurements such as head and abdominal diameters, does about as well as head and trunk 3D reconstructions at predicting weight. However, neither method does well enough to meet our initial goal of better than 200 gms/kg. The most likely explanation seems to be that although we are acquiring a large amount of information, we are not processing it in an "intelligent" manner. Computer plots of the fetal reconstructions show that because of our repeatability errors (see C3 above) and because of fundamental limitations of ultrasound, we are not able to obtain complete and accurate 3D fetal surface points. Although a human looking at the data can easily infer the fetal shape the computer is unable to do this because it has no notion of the shape of a fetus. Thus this study merely justifies more than ever the need for a model based approach.

The research that is currently underway includes the following:

1. Continuation of the clinical study on live term babies since 31 patients is not enough to make definite conclusions. Also since this data will remain available, the evaluation of a model based system should be much easier in the future.

2. In vitro testing of the accuracy of the system on hearts, in collaboration with the Division of Cardiology. Methods of measuring true left ventricular volume are being perfected along with the collection of data. We expect (and have already partially shown on balloons and kidneys) that the current system will do fairly well on hearts, which are much simpler than fetuses. However, even if accurate heart volumes are achieved a model directed approach will be necessary in order to reduce the extreme cumbersomeness of the present system. (It now takes about 2 hours to manually outline scans and to produce a volume).

3. A literature survey is being conducted in the areas of computer vision and 3D modelling in order to determine what methods have been used in the past to represent 3D shapes. The most promising method at present seems to be generalized cylinders, as proposed by Tom Binford at Stanford. The cylinders would probably be represented as uniform cubic splines, as suggested by Uri Shani at Rochester. An additional

requirement will be shape constraints so that the class of all hearts or fetuses can be parameterized such that the given data determines an instance of the generic model.

D. Publications

Brinkley, J.F., Moritz, W.E., Baker, D.W., "Ultrasonic Three-Dimensional Imaging and Volume From a Series of Arbitrary Sector Scans", Ultrasound in Medicine and Biology, vol 4, pp 317-327.

Brinkley, J.F., McCallum, W.D., Daigle, R.E., "A Distributed Computer System for Fetal Weight Determination", Proceedings of the 24th Annual Meeting of the American Institute of Ultrasound in Medicine, Montreal, August 27-31, 1979, p 113.

Brinkley, J.F., McCallum, W.D., Muramatsu, S., "Three-Dimensional Display of Fetus, Placenta and Uterus using an Ultrasonic Computer Modelling System, Proceedings of the 25th Annual Convention of the American Institute of Ultrasound in Medicine, New Orleans, Sept 15-19, 1980.

Brinkley, J.F., Muramatsu, S., McCallum, W.D., "Fetal Weight from Three-Dimensional Ultrasonic Data", submitted to the 26th Annual Convention of the American Institute of Ultrasound in Medicine, San Francisco, 1981.

McCallum, W.D., Brinkley, J.F., "Estimation of Fetal Weight from Ultrasonic Measurements", American Journal of Obstetrics and Gynecology, 133:2, pp.195-200, Jan. 1979.

E. Funding Status

1) "Ultrasonic Measurement of Fetal Volume and Weight"
Principal Investigator: W.D. McCallum, M.D.
Assistant Professor
Department of Obstetrics and Gynecology
Stanford University
Funding agency: National Institute of Child Health and Human Development
Number: 1-R01 HD12327-01
Total term and direct cost: 7/1/79-6/30/81, $111,823
Current funding period: 7/1/79-6/30/80, $60,423

2) "Ultrasonic Three-dimensional Organ Modelling", individual postdoctoral fellowship.
Fellow: James F. Brinkley
Sponsor: W.D. McCallum
Funding Agency: National Institute of General Medical Sciences
Number: 1 F32 GM08092-01
Total term and direct cost: 7/1/81-6/30/84 (3 years) $65,452 (stipend)
Current funding from this fellowship: none

II. INTERACTIONS WITH SUMEX-AIM RESOURCE

A. Collaborations

We are collaborating more with medical people than anyone else. The project is located in the Obstetrics Department at Stanford where W.D. McCallum manages the ultrasound patients. We have also been collaborating with Dr. Richard Popp in the Division of Cardiology at Stanford.

B. Sharing and Interactions with SUMEX Projects

Mostly personal contacts with the Heuristic Programming Project and MYCIN project at Stanford. The message facilities of SUMEX have been especially useful for maintaining these contacts. Since the first phase of the project is now essentially complete we have been interacting more with other SUMEX projects in order to develop the AI ideas.

C. Resource Management

In general SUMEX has been a very usable system, and the staff has been very helpful. The only complaint is that it is impossible to get anything done in the afternoons since we always get bumped.

III. Research Plans

A. Project Goals and Plans

As mentioned in Part I we are implementing this system in phases, each phase requiring use of more sophisticated artificial intelligence techniques. The major phases are as follows (in chronological order:

1. Set up prototype system. Perform engineering tests and clinical evaluations of the ability of the system to predict fetal weight, heart and kidney volume.

As mentioned under Highlights of Research Progress, this first implementation phase is complete, as are the initial engineering evaluations. The clinical study to predict fetal weight should be complete at the end of June when the current grant runs out. Heart studies are just beginning with the Cardiology Division and kidney studies are expected to begin soon with the Department of Radiology, in collaboration with Dr. Barbara Carroll. However, our initial patient studies have demonstrated the basic limitations of the system, which are inadequate models and difficulty of use. From a medical point of view the next phases will be attempts to remove these limitations.

2. Explore other methods for geometric modelling, AI techniques of goal directed problem solving.

In order to develop adequate models and control strategy it will be necessary to examine other AI methods of generating models and using them to guide problem solving programs. This phase of the research is now under way. For this aspect of our research the SUMEX-AIM community has been especially useful.

3. Develop program, as outlined in the introduction, with several limitations:

- Only a simple organ will be modelled at first, ie not the entire fetus including limbs.

- The computer will still request certain scans to be retrieved from the video disk but the operator will outline them with the light pen. Since ultrasound image quality is improving so rapidly it makes sense to wait as long as possible before attempting automated border recognition. The models and control strategies developed during this phase should be useful when actual border recognition is attempted however.

4. Extend the technique to more irregular objects structured models will be developed so that the fetal limbs can be included.

5. Add image processing hardware, develop automated border recognition software.

The models developed in the last two phases will be used to guide the border recognition process. As these phases are implemented they will continue to be tested against the clinical data acquired and stored on floppy disk by the data acquisition system. In this way we can develop new ideas while continually upgrading the clinical utility of the system.

B. Justification for Continued Use of SUMEX

The goals of this project seem to be compatible with the general goals of SUMEX, ie to develop the uses of artificial intelligence in medicine. The problem of three-dimensional modelling is a very general one which is probably at the very heart of our ability to see. By developing a medical imaging system that models the way clinicians approach a patient we should not only develop a useful clinical tool but also explore some very fundamental problems in AI.

C. Need for Resources

1. SUMEX Resources

Our current share of the SUMEX resources is adequate.

2. Other Resources

Judging from our present experience it appears that SUMEX could not handle the amount of data required for image processing on digitized ultrasound scans. This is one of the main reasons we are proposing a distributed system in which SUMEX only directs a smaller machine to do the actual number crunching. It is also one of the reasons we are postponing direct digitization until later. As microprocessors become more powerful they will be capable of acting as slaves to an intelligent SUMEX program. The AI program will direct the image processing functions of the micro so that the data is processed in an intelligent way, but SUMEX will only see the results of that processing, not the actual data. We will thus need to keep track of developments in microcomputers so that we can develop this kind of distributed system.

D. Recommendations

Since we are planning to develop a distributed system we would hope to see these kind of systems being developed by the SUMEX resource. Projects that would be of direct interest are networks (such as ETHERNET), personal computer stations, graphics displays, etc.

II.A.3.3 DECIDER Project: The Psychology of Expert Judgment

The Psychology of Expert Judgment

Eric J. Johnson
Department of Psychology
Stanford University

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

This project focuses on the psychology of expert judgment under uncertainty. There are two separate reasons for pursuing this topic. First, there seems to be a conflict in the view of expertise in two different disciplines. Experts have started to receive much attention in artificial intelligence and cognitive science. The reason for this interest, by and large, is their apparent success in various task domains. Behavioral Decision Theory, in contrast, has called the existence of expert judgment in question, and supplies a number of empirical studies supporting this conclusion. Further exploration of these two conflicting views seems warranted. The current project concentrates on further empirical research and the development of a conceptual analysis of the skills experts bring to judgment tasks.

Second, existing models of novice choice behavior may be successfully extended to account for expert behavior. These models must incorporate domain expertise in order to account for the observed differences between novice and expert subjects in empirical research. Such models can help provide the basis of user interfaces into expert systems, serving as user models aiding in the construction of queries and instructions to the user.

B. Medical Relevance

The initial empirical investigations have focussed on the selection of medical personnel, in particular, the selection of housestaff for post-graduate training. Further extensions are planned using medical diagnosis. Thus, while the study of expert judgment has obvious implications for medicine, this research also uses medicine as its substantive focus.

II. INTERACTION WITH SUMEX-AIM RESOURCES

My connection with the SUMEX resource has encouraged contact with a number of members of the SUMEX and Stanford community. Although my office is located in a building next to the Computer Science Dept., the use of SUMEX for mail has greatly facilitated my interaction with several members of that community.

This interaction has included discussions with members of the MYCIN group and resulted in a talk discussing this work to the Stanford Heuristic Programming Project luncheon meetings.

III. FUTURE PLANS

I am currently preparing a grant proposal based on the initial results of this work. I anticipate that empirical studies will be conducted in the next year. Currently, various medical diagnosis tasks are being considered. The emphasis will be on the relationship between the decision strategies used by the expert judge and that judge's accuracy. I also anticipate that SUMEX will allow access to expert performance systems. One possibility is the comparison of these systems to simple baseline and actuarial models. For these reasons, SUMEX provides a unique resource for this research.

II.A.4 Pilot AIM Projects

The following are descriptions of the informal pilot projects currently using the AIM portion of the SUMEX-AIM resource or the Rutgers-AIM resource pending funding, and full review and authorization.

II.A.4.1 AI-COAG: Coagulation Expert Project

COAGULATION EXPERT Project

Dr. D.A.B. Lindberg
School of Medicine
University of Missouri-Columbia

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

Preliminary experiment to form a clinical consultant program based on a formal representation of medical knowledge of the Blood Coagulation Expert.

B. Medical Relevance and Collaboration

Experts in clotting are few and tend to be based at University hospitals or large tertiary care facilities. It would be extremely helpful if this knowledge could be made available to physicians via an automated system.

Relevance of such a proposed system would be with respect to diagnosis, management, and continuing medical education.

The team at UMC consists of the following individuals:

Lamont Gaston, M.D.
Lawrence C. Kingsland III, Ph.D.
Donald A. B. Lindberg, M.D.
Johannes Yesus, M.D.
Anthony Vanker, Ph.D.

Dr. Gaston is a consulting hematologist, Director of the Coagulation Laboratory and Dr. Yesus is Director of the Blood Bank. Dr. Lindberg is Director of the Information Science Group; Dr. Vanker is a Postdoctoral Fellow whose Ph.D. is in Physiology; and Dr. Kingsland is a Postdoctoral Fellow whose Ph.D. is in Electrical Engineering.

Expertise in the field as well as clinical laboratory and patient records are being provided by UMC to build and test the consultant. In addition, testing of the initial program and consultation has been obtained from Heinz Joist, M.D., Director, Coagulation Laboratory, St. Louis University Hospitals.

A formal research proposal was submitted to NIH on October 27, 1980, based on the studies performed on SUMEX and at UMC.

C. Highlights of Research Progress

Accomplishments:

Use of UNITS/AGE: an initial model has been created on SUMEX.

Experimental use of EMYCIN: a feasibility test with a textbook level consultant model has been created on SUMEX.

Use of local LSI-11: In addition the initial knowledge base has been assembled into a simpler (but operational) system on a DEC LSI-11 using RT-11 and BASIC.

We have selected a strategy for development. This is to begin with the interpretation of clinical laboratory tests: first the Coagulation Screen (of 6 tests), then the Partial Coagulation Screen (of 3 tests), then the individual determinations. Laboratory and clinical features will be taken into account, and other features will be added.

Research in Progress:

We have tested the initial model against actual clinical records for 315 patients. This is partly as a validation of the work done, and partly as a means to bring to our attention the unusual circumstances and unforeseen problems which we know will be present. That is, we have allowed for all feasible patterns of results but (probably) have not yet allowed for all the surrounding clinical circumstances. Initial results: 80% of model responses are appropriate as is. 20% are inappropriate as is, because the model does not yet allow for patient clinical history and certain special assays. We are now constructing a clinical history-taking program.

D. List of Relevant Publications

Lindberg, D.A.B., Gaston, L.W., Kingsland, L.C. III, Vanker, A.D., AI/COAG: A Knowledge Based System for Consultation about Human Hemostasis Disorders: Progress Report. Submitted to Fifth Annual Symposium on Computer Applications in Medical Care, November 1-4, 1981, Washington, D.C.

E. Funding Support

This preliminary research phase was supported from two sources.

1. USPHS Grant No. T15 LM 07006, "Training Program in Medical Information Science." Full funding is $161,410/year. About $25,000/year is being devoted to this project.

2. USPHS Grant No. HS 02569, "Health Care Technology Center." Current funding is $500,000/year. About $12,000/year was devoted to this project.

A project research proposal has been submitted jointly to the National Library of Medicine and the Blood Resources Program of the National Institutes of Health.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations and Program Disseminations via SUMEX

Dr. Vanker will give an oral presentation of our work at the Spring meeting of Trainees and Directors, NLM Training Programs, in May at Washington, D.C. We have also given individual demonstrations of our models to visiting scientists (including some from Japan) at UMC.

B. Sharing and Interactions with Other SUMEX-AIM Projects

In February, 1980, David Goldman, a medical student at UMC and former pre-doctoral fellow in the Information Science Group, spent a week at Stanford University becoming acquainted with the various artificial intelligence systems in development at SUMEX. In fact, with the help of members of the SUMEX-AIM community, he was able to implement a simple workable coagulation model in EMYCIN.

In March, 1980, Dr. Ueno attended a workshop on AGE at Stanford. Through this workshop he was able to learn a great deal that was directly applicable to our work. He also obtained a better understanding of the UNITS package and how it might be used to interface with AGE. All of us in the study group attended the AIM-tutorial at Stanford in August 1980.

Since the AI systems in which we are interested in are in some stage of implementation on the SUMEX computer and since partial documentation does exist for these systems, we have been able to learn a great deal on our own by an interactive, trial-and-error method. Of course we have had many questions and we have been able to obtain prompt and helpful information from various members of the SUMEX-AIM community via the network electronic message system.

Our decision to make the main line of our development based on the LSI-11 microprocessors is possible only because of the existence of SUMEX and the larger central system as back-up to us should the problem exceed our local capacity.

C. Critique of Resource Management

We have found the people at SUMEX to be uniformly helpful and more than willing to aid us in our attempts to understand the various aspects of artificial intelligence (AI) in medicine. Both Mr. Goldman and Dr. Ueno were delighted with their experiences at Stanford and commented on the willingness of otherwise very busy people to help them with their problems.

One of the drawbacks of SUMEX is that quite often the interaction is slow. There are days when we must wait up to several minutes between exchanges between our terminal and SUMEX. This is apparently due to a high average load on SUMEX at the time. We have had no other problems with the resources at SUMEX and we feel the management has done a good job.

It is clear that truly collaborative research with SUMEX (as opposed to research which merely draws on SUMEX computational resources) requires the designation of a collaborating computer scientist at Stanford. Collaboration of this sort can arise out of personal contact through events such as the AIM Workshop, through sabbaticals, and scientific meetings. In addition, collaboration typically involves capturing the interest of graduate students, and must provide these students with problems which are suitable for dissertation research. These things have not happened as yet in the case of the AI/COAG project but we look forward in the future to this possibility. In any event, this necessary personal component in remote collaboration teaches us that the training, education, and dissemination aspects of resources such as SUMEX are extremely important and cannot be separated from the research elements.

III. RESEARCH PLANS

A. Project Goals and Plans

Plans for the Coming Year:

1. Continue assembling the knowledge data base, with emphasis on documentation of the primary literature sources of the KS's.

2. Continue to study UNITS/AGE and EMYCIN.

3. Continue comparison of the two potentially complex models vs. the microprocessor version.

Long Range Plans:

Our long range plans are to develop the consultation system with Knowledge Sources in the following areas: interpretation of laboratory screening tests, hemostasis patient history, management of oral anticoagulants, diagnosis of hemophilia/von Willebrand's disease, platelet function studies, diagnosis and management of disseminated intravascular coagulation, and pre-operative hemostasis studies.

B. Justification and Requirements for Continued SUMEX Use

As our knowledge base grows, the capabilities of UNITS/AGE will become increasingly more important to us. The UNITS package has built-in means of dealing with large amounts of knowledge in a hierarchical fashion. AGE is a knowledge-based program design to build other knowledge-based programs.

An ancillary, but still important, objective of our work in AI in medicine is to learn about the strengths and weaknesses of the particular AI programming systems in use and development in order to understand better how knowledge can be stored and manipulated. This understanding, in itself important, may then be applied in the design of simpler, but perhaps more accessible programs which can be implemented on micro- or mini-computers.

II.A.4.2 DATA Project

DATA: Data Analysis Tutorial and Advisory System

Asst. Prof. Ruven E. Brooks
Department of Psychiatry and Behavioral Sciences
University of Texas Medical Branch

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

1. The Problem of Biomedical Data Analysis Skills

The lack of sophistication of most biomedical researchers in data analysis is a pervasive constant in the professional functioning of journal editors, funding study groups, and the like. The deficiencies range from the merely regrettable, such as the use of a few too many t-tests or failures to use all applicable analysis techniques, to ones which invalidate entire data analyses, such as failures to substantially meet the statistical assumptions underlying analysis techniques or erroneous interpretation of analysis results. While the source of all of these problems can be covered with the broad label, "inadequate knowledge of statistics," they can, in fact, be traced back to several distinct knowledge deficiencies:

(1) Insufficient knowledge of statistical and probability theory. The type of problems this gives rise to are ones in which valid data have been collected, but the analyses used significantly violate tenets of statistical theory. Typical examples might include performing multiple tests without correction of significance levels ("probability pyramiding") or the use of an improper procedure for estimating missing data.

(2) Faulty data analysis inference/faulty "statistical logic." This phrase is used to describe lack of knowledge about the relationship between questions being asked about the data and the information yielded by analysis techniques. The problems that this gives rise to are ones in which valid data have been collected and statistically acceptable analyses have been performed but in which results from the analysis technique don't match the given scientific question. Typically, these problems manifest themselves as misinterpretations of the results of analyses, for example, using the order of entry of variables into a stepwise regression as an indicator of how strongly the individual variables predict the criterion. Additionally, they can also reveal themselves in inappropriate uses of inferential analyses when exploratory approaches would be more appropriate.

(3) Faulty scientific inference. This is intended to cover situations in which the scientific or experimental questions do

not match the data that was collected, i.e., in which the variables being measured cannot answer the scientific question being asked, regardless of the statistical analysis used. While there are many humorous stories of researchers asking statisticians to make sense out of useless data, there is a fine line dividing errors of this type from legitimate scientific questions as to the correct interpretation to be given to particular measurements.

We assert that deficiencies of the first type can be eliminated by more thorough and wider use of existing tools and techniques, and that deficiencies of the third type must be resolved by better training within individual biomedical disciplines. We will argue, however, that remedying deficiencies of the second type require a new approach.

2. Deficiencies of Existing Approaches

2.1 Improved Graduate Education

Early initial and substantial training in the skills of statistics and data analysis must be a part of any worthwhile graduate program in the biomedical sciences. However, more, better, or longer courses in biomedical statistics do not entirely solve the problem of providing the working biomedical research with data analysis competence in useable form. For most researchers, the design of an experiment and the analysis of the final results together are but a small fraction of their research activities, occupying a relatively small portion of their work time. Hence, their data analysis skills, however good they may have been originally, are certain to have deteriorated significantly by the time they are needed. Moreover, data analysis techniques are continuously evolving, and while this evolution may take place relatively slowly from the statistician's viewpoint, it still may take place rapidly enough to rule out the feasibility of one-time education.

The previous paragraph suggests another possibility, continuing education for researchers, in analogy to continuing education programs for physicians. Indeed, there are short courses and seminars for researchers on both analytic techniques and particular computer packages. However, these courses are not of value to most researchers since they are rarely scheduled to coincide with the completion of a study and the deadline for a final report or journal submission.

2.2 Self-Educational Materials

Given the inadequacy of formal courses, most researchers rely on one of two strategies: The first is to use self-education to recover the skills acquired during graduate education. (This is, of course, presuming that adequate mastery of these skills was obtained in the first place.) Indeed, this can be perfectly adequate for brushing up on one's knowledge of statistical or probability theory, for example, for reviewing the assumptions that must be met to use normal analysis of variance. It is not, however, much use in learning or relearning the process of statistical logic or data analysis inference. This process involves first, selecting a statistical analysis technique that is appropriate both to the properties of the measurements that were made and to the scientific question being asked and, then, interpreting the results of the analysis in terms of this scientific question. Useful tutorial materials in the logic of data analysis must help users relearn or learn this process, rather than just a set of static facts.

Existing materials are not well suited to teaching this process. Consider, first of all, statistics textbooks or sourcebooks. Almost all of these texts are organized by analytic technique or model, not by the structure of the data or by the question to be answered by the data analysis. This organization is not well suited to teaching the process of data analysis inference. Suppose, for example, that the researcher had collected categorical data and wanted to review methods for determining sets of predictors for it; the relevant material would be so scattered through the chapters of most multivariate analysis sourcebooks that the researcher is likely to miss important parts of it.

A form of tutorial material that does not share the drawbacks of the textbook or sourcebook is the case study book which follows the analysis of a complete case from the formulation of the initial data analysis questions or goals through the selection of analytic techniques to the interpretation of final results. While such casebooks could be useful, they are rarely available in the researcher's area of expertise and it is usually difficult to locate the cases that are relevant to the researcher's problem.

Consultants:

Given the inadequacy of sourcebooks and the lack of sufficient case books, most biomedical researchers today use statistical consultants as their source of knowledge for performing data analyses. As convenient as this may be for the researcher who has a good consultant available, it is by no means desirable as a permanent solution to the problem. Such consultants are invariably overburdened and pressed for time. They, therefore, are forced to structure their task into one of learning, as quickly as possible, the minimum knowledge needed to understand the user's research problem and then telling the user what analyses to perform and how to interpret them. Teaching the researcher the rationale behind the analysis selection and interpretation frequently becomes an option to be indulged in if there's any time left over. Thus, interaction with a consultant rarely results in much of an improvement of the biomedical researcher's own skills.

From the consultant's perspective, current practice also leaves much to be desired. After finally learning enough about the research domain to understand the issues, many problems turn out to be, if not routine, at least ordinary enough to present little challenge while others turn out to be hopeless, no matter how much creativity the consultant supplies. Where, ideally, the consultant should play the role of expert confirming the applicability of standard practice or creating a new analysis for an unusual case, in current practice, the consultant's role is frequently just to compensate for deficiencies in the skills a competent researcher should be expected to display.

3. Proposed Research

The primary goal of this research is the construction of a computer system to teach and advise users on how to statistically analyze their data. The intended user population is biomedical researchers who have a rudimentary knowledge of statistical techniques as might be acquired through one or, even, several graduate courses, but, who, because they do data analysis or experimental design infrequently, require tutoring and guidance. It is not intended to provide complete initial training in statistics, though it could be a useful tool in such training.

The basic method of operation of the system would be to guide users through case studies of the logical process of selecting data analysis methods and interpreting their results, based on the nature of the measurements and on the experimental questions. The guidance would take the form of both feedback to specific user decisions and a general question answering capability. This latter capability could be used to generate explanations on topics ranging from the reasons for analysis selection and interpretation in the specific case at hand to background questions about how analyses are performed, what the validity criteria are for interpreting them, etc. The amount and level of detail in these explanations would be controlled by a user model based on the individual user's sequencing and on the details of the case.

The processing of a case will take place in four phases. In the variable description phase, the user will be shown a description of what was measured and asked to characterize it with terms such as metrical versus categorical, maximum-minimum values, etc. In the question clarification phase, the user will be given a scientific question to be answered in the study and asked to recast it in terms of questions about the variables that were actually measured. For the analysis selection phase, the user will be asked to select analyses and will be given feedback in terms of reasons for a selection to the quality of the selections. Finally, in the analysis interpretation phase, the user will be shown the results of an analysis for different sets of sample data, and asked to give a correct interpretation.

The system will have three main operating modes. In the stored case mode, the properties of the variables, the data analysis questions, and the data analysis technique will all be stored in fixed form. In the update mode, users will be allowed to alter the properties of the variables and the analysis techniques and to observe the effects on question formulation, analysis selection, and analysis interpretation. Finally, in new case mode, the user will enter all of the information, starting with the scientific question. While this mode will not permit checking whether the data analysis questions are consistent with the scientific questions, it will still be useful as a tool for permitting users to explore potential experimental designs.

4. Expected Benefits of this Research

The proposed tutorial system would have a number of significant advantages over current approaches. First, because it is organized around the case, rather than around the properties of individual analysis techniques, it will do a better job of conveying the relationships between scientific questions, measurements, analysis techniques, and interpretation. Further, since the presented cases can be tailored to fit the problem the researcher has at hand, there is less likelihood that significant and relevant information will be overlooked. Additionally, since the component that generates the example cases is separate from the one that generates the text to describe them, it is easy to present the cases in the language of the researcher's specialty.

A second advantage of the proposed system is that it could be used by researchers to explore potential experimental designs before actually running an experiment. Though researchers are frequently advised to talk to a statistician before beginning a study, the prospect of attempting to schedule several meetings with a busy consultant is frequently enough to discourage researchers from following this advice. A learning tool, such as the one proposed here should prove conducive to better designing of studies.

These two direct advantages have the potential to combine and produce an indirect benefit: greater productivity on the part of statistical consultants. The argument can be made that researchers should always go to a consultant, regardless of their own level of competence, if only to get a second opinion from someone equally competent. Given the validity of this argument, the more the researcher knows, the more useful the consultation is likely to be with enhanced opportunities to produce the creative or optional analysis, rather than the merely adequate one.

C. Highlights of Research Progress

During this past year, proposals were submitted to the National Science Foundation and to the National Institutes of Health; both were declined. The critiques from both sets of reviews shared a common theme, that the construction of a computer system to give advice on statistical analysis was not a desirable goal, whether or not it was technically feasible. The basis for this position was that any system which increased an investigator's reliance on a computer system, rather than on his or her own intellectual resources, would lead to a decline in research quality. To quote one of the more vituperative NSF reviewers:

> ...If successful, the system purportedly would reduce the need for research assistants who interface with the computer systems, as well as some portions of what statistical consultants do and, I greatly fear, further reduce the amount of thinking that goes into analysis.
>
> ...More importantly, however, the thrust of the proposal is all wrong. It suggests the use of computers in precisely the way they are most inappropriate, that is, trying to substitute computer generated rules for human thought. While a portion of the program may be of some use, the main thrust of the proposal is to substitute advice based on a set of computer generated questions and

> answers for the investigator's own knowledge. Just as the existence of the computer software packages have enabled people who don't understand statistical techniques to use them by reducing the cost of computation, so, I think, this program would give people a false sense of security and lead them to think they were using correctly techniques they do not understand. The effect of this program, if it were successful, would be to further reduce the thought that goes into designing and analyzing data, and make research even more mindless than it all too frequently is now.

Since the nature of this criticism left little hope of ever obtaining funding for the project as originally conceived, a major re-orientation of the project was undertaken, shifting from a purely advisory system towards a tutorial system. This shift has had a number of consequences, the most trivial of which being the development of a new acronym (DATA system = Data Analysis Tutorial and Advisory system). More important, it has caused a shift in the intellectual focus of the project from being primarily one of the application of existing software to a new domain, to one that touches on two issues of more fundamental A.I. interests, methods for providing explanation and methods for problem generation in intelligent CAI.

D. List of Relevant Publications

None.

E. Funding Support

None.

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaborations via SUMEX

B. Sharing and Interaction with Other SUMEX-AIM Projects

In August, 1980, I attended the AIM workshop at Stanford. While in Palo Alto, I visited Tom Moran at Xerox PARC who gave me a manual for the Interactive Data Language, an Interlisp based statistical package developed by Ron Kaplan and Beau Shiel. Since this system has useful potential for both the DATA project and the RX project, I inquired whether it was available at SUMEX and learned that, somewhat surprisingly, SUMEX projects were unaware of its existence. It has since been transferred down the hill and is now being used by the RX project and will be available for use in the DATA project.

C. Critique of Resource Management

As this project has not involved significant utilization of machine resources over the past year, no comments in this area are appropriate.

III. Research Plans

A. Project Goals and Plans

The major short-term goal for the next year will be to complete the first-level software design for the entire system and to begin construction of the problem generation component. This component, which generates case descriptions of statistical analysis problems, is, by itself, an operational system since the case descriptions can be evaluated and used independently of the rest of the system.

B. Justification for Continued SUMEX Use

SUMEX continues to be the only nationally available resource that is oriented towards support of the remote (as versus on-site) artificial intelligence in medicine researcher. While I am located in the nation's sixth largest medical school, I am not located near any suitable research facilities; the artificial intelligence group at Austin is more than 200 miles away, and my local environment has neither the hardware nor the software support for any kind of artificial intelligence work nor is there any possibility of acquiring any before 1983.

C. Needs and Plans for Other Resources

The Medical Computer Science Dept. at the University of Texas Health Science Center in Dallas is exploring the possibility of establishing a facility for A.I. research. Should their efforts prove fruitful, it would provide a feasible alternative to SUMEX, though not for at least 2-3 years.

D. Recommendations for Future Resource Development

If the efforts with single-user systems prove fruitful at Stanford, it will be important to develop plans for software licensing and support at other sites. Given that the true annual cost of an on-site programmer to maintain a personal computer is a substantial share of the one-time cost of the machine, there will still be considerable need for SUMEX to provide support even if it becomes possible for each research group to have their own hardware.

II.A.4.3 EXCHANGE Project

Japan/U.S. Medical Consultation System Exchange

Drs. S. Kaihara, T. Koyama,
and F. Mizoguchi
Hospital Computer Center
University of Tokyo Hospital

I. SUMMARY OF RESEARCH PROGRAM

A. Project Rationale

Our goal is two-fold, namely the international comparison of medical knowledge formalized in the consultation systems in some clinical areas, and secondly, the trial of cooperative development of computer programs in the international settings. Both objectives can be achieved only by using such facilities as SUMEX-AIM which can be accessed internationally by the collaborating researchers.

B. Medical Relevance and Collaboration

Due to rapid increase of medical knowledge required for daily practice of medicine, the interest in medical consultation systems is increasing. Our project involves two most prominent expert physicians, Prof. Shigenobu Nagataki, Professor of Medicine at Nagasaki University and Prof. Yoshiyuki Suzuki, Professor of Pediatrics at University of Tokyo, who are both enthusiastic for the development of consultation systems in their field.

C. Highlights of Research Progress

Three projects are in progress:

1) Comparison of medical expert knowledge formalized in the system at international settings. Two systems are now being developed using the EXPERT system developed at Rutgers University. The first one is the consultation system of thyroid diseases led by Prof. Nagataki. The core of the system has been completed and the knowledge will be compared with the American expert knowledge in the near future.

2) Development of practical consultation system using "AGE". Our experience showed that the diagnosis obtained by deduction covers only a small part of physicians' interests. They often express a need for such functions as data error checking, strategies for obtaining data, evaluation of missing data on the diagnosis, reference data related to the diagnosis, etc. A consultation system which has the above features is now being developed, using "AGE" which was developed at Stanford University. Preliminary results in the consultation of inborn errors of metabolism showed encouraging results.

3) Development of MECS-AI. A general purpose system named as MECS-AI is under development. This system is designed to assist in constructing knowledge-based inference systems with time-oriented data treatments.

D. List of Relevant Publications

None yet.

E. Funding Support

i) The Japan-US Cooperative Science Program (grant)
Shigekoto Kaihara, Hospital Computer Center, University of Tokyo
The Japan Society for the Promotion of Science
(no number)
March 1979 to February 28, 1981 5,280,000 Japanese yen

ii) Grant in Aid for Scientific Research (grant)
Shigekoto Kaihara, Hospital Computer Center, University of Tokyo
The Japan Society for the Promotion of Science
549009
April 1980 to March 1983 8,736,000 Japanese yen

II. INTERACTIONS WITH THE SUMEX-AIM RESOURCE

A. Medical Collaboration and Program Dissemination via SUMEX

Dr. Lindberg's group at University of Missouri is collaborating with us for the evaluation of consultation programs.

Dr. Kulikowski's group at Rutgers University is also collaborating by supplying EXPERT and evaluating developed systems.

B. Sharing and Interactions with Other SUMEX-AIM Projects

We are collaborating with the "AGE" developing project at Stanford University, which supplies us with the newest version of AGE. Our group also attended the AIM workshop held at Stanford in August, 1980.

C. Critique of Resource Management

We have found the staff professional and helpful. We have had almost no troubles in using SUMEX facilities from Tokyo, Japan.

III. RESEARCH PLANS

A. Project Goals and Plans

Project goals are two, namely 1) international comparison of medical knowledge formalized in the consultation system in clinical fields; 2) international collaboration for development of more practical medical consultation systems using AGE or MECS-AI. The procedures described in Section I.C. will be continued.

B. Needs and Plans for Other Computing Resources beyond SUMEX-AIM

We are using computer facilities located in Japan whenever possible to reduce the cost of international data communication. This will be continued.

C. Recommendations for Future Community and Resource Development

Development of systems which can be transferred to smaller machines such as VAX is desirable in our environment, for there is only one DEC 20 in the large city of Tokyo.

II.A.4.4 MELANOMA Project

Decision Support for Clinical Melanoma

Drs. M.S. Blois and Hyo Kim
Section on Medical Information Science
University of California Medical Center
University of California at San Francisco

We have not made significant progress on our project of "Decision Support for Clinic Melanoma". Due to the low priority we have and the large size of EMYCIN program we had only limited access to the computer. However we have become familiarized with the system and we hope that we will be able to make slow but steady progress in future.